2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Distributionally Constrained Convex Duality Optimal Control (DC-CDOC) Subject to Different Forms of Constraining the Terminal State of Nonlinear Stochastic Systems

Ali Pakniyat, *Member, IEEE*

Abstract— **This article explores different methods of constraining the terminal state of nonlinear stochastic systems and presents the results of the Distributionally Constrained Convex Duality Optimal Control (DC-CDOC) for each case. Specifically, we consider: (a) almost surely equality constraints on the terminal state, (b) constraints on the expected value of the terminal state, and (c) constraints on the probability distributions of the terminal state, both (i) under the total probability measure and (ii) under all conditional probability measures. For each case, the associated optimal control problem is formulated as a convex linear program on the space of Radon measures, and by exploiting the duality relations between the space of measures and that of continuous functions, we derive the optimality conditions in the form of an optimization problem over the space of differentiable functions constrained to a Hamilton-Jacobi (HJ) inequality and, in some of these cases, a terminal value inequality. An iterative algorithm is also proposed for identifying the value function in the studied cases.**

I. Introduction

This paper addresses finite-horizon optimal control problems for continuous time nonlinear stochastic systems, where the control objective is to steer the state from an initial condition to a desired terminal probability distribution with known statistics. In the literature, problems of this type have only appeared for special subclasses of systems. More precisely, the majority of studies assume linearity of the dynamics and a quadratic form for the cost, and Gaussian forms for the desired distribution. The associated results are presented for both infinite time horizon problems [1]–[4] and finite time horizon problems in both continuous time and discrete time settings [5]–[13]. The accommodation of input constraints is considered in [9], and convex relaxations for linear systems subject to chance constraints are studied in [11], [12]. Nonlinear extensions of the probability distribution assignment have been presented only for special subclasses, including feedback-linearizable systems [14], and gradient flow systems [15], [16], or through approximating methods such as sequences of iterative linearizations [17] and differential dynamic programming approximations [18].

In past work of the author [19]–[22], two distinct novel viewpoints are presented which apply to a wide range of nonlinear stochastic systems. The first approach [19]–[21] is a perpetual renewal of a constraint on the expected value of the terminal state whose associated optimality conditions are established in the form of the Terminally Constrained Stochastic Minimum Principle (TC-SMP). While the TC-SMP applies to both linear and nonlinear systems, a major limitation of the theory is that it can only impose constraints on the first moment (expected value) of the terminal state. By reimposing this constraint under conditional expectations at all filtrations at future times, the TC-SMP is very successful in steering the state towards the desired location for almost all sample paths, however, this is achieved at the expense of increased control effort close to the final time. The second approach [21], [22] is an extension of the covariance control problems for general nonlinear systems with nonlinear costs and general desired probability distributions (i.e., non-necessarily Gaussian desired distributions). This, in particular, calls for a change of viewpoint from the study of *sample paths* (where the terminal state distribution is a statistical byproduct of the investigation) to the study of the so-called *occupation measures* in which the description of the terminal distribution constraint is inherently natural to the representation. Accordingly, the assignment of probability distributions to the terminal state is reformulated as a convex linear program over the space of measures and the associated optimality conditions are established by invoking the duality relationships between the space of measures and that of continuous functions.

The convex duality method for optimal control problems was initiated by Vinter and Lewis [23], [24] for deterministic control systems and, later, by Fleming and Vermes for piecewise deterministic [25] and stochastic [26] processes. The fundamental idea of this approach is the introduction of a weak formulation that embeds the original (strong) problem into a convex linear program over the space of Radon measures. Upon establishing the equivalence of the two problems, new necessary and sufficient optimality conditions are obtained by invoking the Fenchel-Rockafellar duality theorem. This approach is particularly useful in characterization of optimal policies in certain desirable classes of controls by investigating the extreme points of the set of Hamilton-Jacobi problems (see e.g. [27]–[30]). For convex duality based numerical algorithms for deterministic continuous systems, one can refer, e.g., to [31]–[33].

In this paper, we elaborate on the Distributionally Constrained Convex Duality Optimal Control (DC-CDOC) of [21], [22] to accommodate various other forms of constraints on the terminal state of nonlinear stochastic systems. To be specific, we consider (a) almost surely equality constraints on the terminal state, (b) constraints on the expected value

A. Pakniyat is with the department of Mechanical Engineering, University of Alabama, Tuscaloosa, AL, USA (e-mail: apakniyat@ua.edu).

979-8-3503-1632-2/24/$31.00 ©2024 IEEE

of the terminal state, and (c) constraints on the probability distributions of the terminal state, both (i) under the total probability measure, and (ii) under all conditional probability measures, and we present the optimal conditions of DC-CDOC associated with each of these cases.

The organization of the paper is as follows. Section II presents the underlying dynamics and costs for class of stochastic optimal control problems, which are subject to constraints on their terminal state in the forms of the constraints presented in Section III. The associated measure theoretic formulation of the constrained optimal control problems are presented in Section IV and the associated optimality conditions of DC-CDOC are presented. Section V provides a conceptual algorithm to establish the value function of the studied cases. Concluding remarks and future research directions are presented in Section VI.

II. The Underlying Dynamics and Costs

Let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t_0}^{t_f}, \mathsf{P}\right)$ be a filtered probability space with $\{\mathcal{F}_t\}_{t_0}^{t_f}$ an increasing family of sub σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_{t_0}$ contains all the P-null sets, and $\mathcal{F}_{t_f} = \mathcal{F}$ for the fixed terminal time $t_f < \infty$. Consider a nonlinear stochastic systems governed by the controlled Itô differential equation

$$\mathsf{d}x_s = f\left(s, x_s, u_s\right) \mathsf{d}s + g\left(s, x_s\right) \mathsf{d}w_s, \tag{1}$$

where $x_s \in \mathbb{R}^n$, $u_s \in U \subset \mathbb{R}^m$, and $w_s \in \mathbb{R}^d$ are, respectively, the values of the state, the input, and the realization of a standard Wiener process at time $s \in [t_0, t_f]$. The input value set U is assumed to be convex and compact and the functions f and g are considered to be Lipschitz functions over, respectively, $[t_0, t_f] \times \mathbb{R}^n \times U$ and $[t_0, t_f] \times \mathbb{R}^n$ with linearly bounded growth rates.

Let $u_{[t_0,t_f]} := \{u_s : t_0 \le s \le t_f\}$ denote a nonanticipative, U-valued, input process such that $u_s \in U$ is progressively measurable with respect to $\mathcal{F}_s$ for all $s \in [t_0, t_f]$. We denote by $\mathcal{U}_{[t_0,t_f]}$ the set of all such inputs.

For a given initial condition $x_0 \in \mathbb{R}^n$ at $t = t_0$, we associate to each $u_{[t_0,t_f]} \in \mathcal{U}_{[t_0,t_f]}$ a total cost

$$J\left(t_0, x_0, u_{[t_0,t_f]}\right) = \mathbb{E}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}\left[\int_{t_0}^{t_f} \ell\left(x_s, u_s\right)\mathsf{d}s + L\left(x_{t_f}\right)\right] \tag{2}$$

where ℓ and L are continuous functions with polynomial growth and $\mathbb{E}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}[\bullet] := \mathbb{E}[\bullet \,|\, \mathcal{F}_{t_0}; u_{[t_0,t_f]}]$ denotes the conditional expectation under the filtration $\mathcal{F}_{t_0}$ and given the input $u_{[t_0,t_f]}$.

III. Constraining the Terminal State

Let $m_d \in \mathbb{R}^n$ denote a desired value for the state of the system to which we would like the controller to bring the state of the system at the terminal time t_f. Let

$$x_\tau^{\mathcal{F}_t} \equiv x_\tau = x_t + \int_t^\tau f\left(s, x_s, u_s\right) \mathsf{d}s + \int_t^\tau g\left(s, x_s\right) \mathsf{d}w_s \tag{3}$$

denote the random variable representing the state at a future time $\tau \in [t, t_f]$ given the filtration $\mathcal{F}_t$ and under the input $u_{[t_0,t_f]}$[1]. It shall be remarked that under the assumption of full observation of the state $x_t^{\mathcal{F}_t} = x_t$ is a known (deterministic) value whereas $x_\tau^{\mathcal{F}_t}$ is a random variable for $\tau > t$.

Because direct constraints on the terminal state such as $x_{t_f}^{\mathcal{F}_t} = m_d$ are non-causal (hence, cannot be achieved by causal inputs), we restrict attention to causal constraining methods. In particular, the following methods of constraining the terminal state are studied in this paper.

A. Almost Surely Satisfactions

The strongest form of constraining the control problem is to required the terminal constraint to match the desired value almost surely, which can be interpreted in the following two ways.

1) Almost Surely Satisfaction in Total Probability: The constrained optimal control problem becomes the minimization of the cost (2) subject to the dynamics (1) and the constraint $x_{t_f}^{\mathcal{F}_{t_0}} \overset{a.s.}{=\!=} m_d$, which can equivalently be written as $\mathsf{P}\left(x_{t_f} = m_d\right) = 1$. Since $\mathcal{F}_{t_0}$ contains all the P-null sets, this is equivalent to

$$\mathsf{P}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}\left(x_{t_f} = m_d\right) = 1, \tag{4}$$

where $\mathsf{P}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}(\cdot)$ denotes the conditional probability under the filtration $\mathcal{F}_{t_0}$ of a state-dependent event under the input $u_{[t_0,t_f]}$.

2) Almost Surely Satisfaction in Conditional Probabilities: The constrained problem is formed by imposing the family of constraints

$$\mathsf{P}_{\mathcal{F}_t}^{u_{[t,t_f]}}\left(x_{t_f} = m_d\right) = 1, \qquad t \in [t_0, t_f]. \tag{5}$$

This family of constraints includes, in particular, the constraint (4) but it contains uncountably many other constraints for any $t \neq t_0$.

B. Satisfaction in Expectation

A weaker (more relaxed) method of constraining the problem is to require the expected value of the terminal state to match the desired value. Similar to the previous case, this requirement can be imposed only on the total expectation, or on the family of conditional expectations.

1) Satisfaction in Total Expectation: The constrained problem is formed by imposing the constraint $\mathbb{E}[x_{t_f} - m_d] = 0$, which is equivalently written as

$$\mathbb{E}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}\left[x_{t_f} - m_d\right] = 0. \tag{6}$$

2) Satisfaction in All Conditional Expectations: The constrained problem is formed by imposing the family of constraints

$$\mathbb{E}_{\mathcal{F}_t}^{u_{[t,t_f]}}\left[x_{t_f} - m_d\right] = 0, \qquad t \in [t_0, t_f]. \tag{7}$$

[1] $u_{[t,\tau]}$ to be more precise because, due to the causality of the system, the evolution of the state within the interval $[t, \tau]$ depends only the restriction of $u_{[t_0,t_f]}$ to the time interval $[t, \tau]$.

C. Satisfaction in Probability Distributions

Another method of constraining the control problem is to require the probability distribution of the terminal state to match, either exactly or approximately, a specific desired probability distribution. These requirements are a generalization to nonlinear systems of the literature on assigning terminal probability distribution to the state linear stochastic systems [1]–[13]. Let p_d be a probability distribution, i.e.,

(i) $\mathsf{p}_d \geq 0$.

(ii) $\int_{\mathbb{R}^n} \mathsf{p}_d(\mathsf{d}x) = 1$,

(iii) $\int_{B_x^1 \cup B_x^2} \mathsf{p}_d(\mathsf{d}x) = \int_{B_x^1} \mathsf{p}_d(\mathsf{d}x) + \int_{B_x^2} \mathsf{p}_d(\mathsf{d}x)$ whenever $B_x^1 \cap B_x^2 = \emptyset$ for any Borel sets $B_x^1, B_x^2 \subset \mathbb{R}^n$.

In particular, for consistency with the requirement that the terminal state reaches (now, in probability distribution) the desired value $m_d \in \mathbb{R}^n$, we require that

$$\int_{\mathbb{R}^n} x\, \mathsf{p}_d(\mathsf{d}x) = m_d \tag{8}$$

We also denote by Σ_d the covariance of the desired probability distribution, i.e.,

$$\int_{\mathbb{R}^n} (x - m_d)(x - m_d)^\top \mathsf{p}_d(\mathsf{d}x) = \Sigma_d \tag{9}$$

Depending on whether the total probability or the family of conditional probabilities are used and whether equality or inequalities are enforced, we present four constraining methods.

1) Exact Assignment of Total Probability Distribution: Following the exact assignment of Gaussian distributions to linear systems in [1]–[13], the problem of assigning a general probability distribution to nonlinear systems is presented in [22] by imposing the constraint

$$\mathsf{P}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}\left(x_{t_f} \in B_x\right) = \int_{B_x} \mathsf{p}_d(\mathsf{d}x), \tag{10}$$

for every Borel set $B_x \in \mathbb{R}^n$.

2) Exact Assignment of the Total Expectation and Containment of the Total Covariance: Following the covariance containment methodology [11], [12], the constrained optimal control problem is defined by imposing the constraints

$$\mathbb{E}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}\left[x_{t_f}\right] = m_d, \tag{11}$$

$$\mathbb{E}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}\left[x_{t_f} x_{t_f}^\top\right] - \mathbb{E}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}\left[x_{t_f}\right] \mathbb{E}_{\mathcal{F}_{t_0}}^{u_{[t_0,t_f]}}\left[x_{t_f}^\top\right] \preccurlyeq \Sigma_d. \tag{12}$$

3) Exact Assignment of Conditional Expectations and Containment of Conditional Covariances: Extending the requirements (11) and (12) to constrain the optimal control problem under all filterations, we impose the constraints defined by

$$\mathbb{E}_{\mathcal{F}_t}^{u_{[t,t_f]}}\left[x_{t_f}\right] = m_d, \tag{13}$$

$$\mathbb{E}_{\mathcal{F}_t}^{u_{[t,t_f]}}\left[x_{t_f} x_{t_f}^\top\right] - \mathbb{E}_{\mathcal{F}_t}^{u_{[t,t_f]}}\left[x_{t_f}\right] \mathbb{E}_{\mathcal{F}_t}^{u_{[t,t_f]}}\left[x_{t_f}^\top\right] \preccurlyeq \Sigma_d. \tag{14}$$

IV. Optimality Conditions

Convex duality relationships between the space of signed measures and that of continuous functions are powerful tools for the identification of the value function and the associated optimality conditions. In particular, with the generalization of the occupations defined in [22] to accommodate general filtrations $\mathcal{F}_t$, we define the *input-state-time occupation measure* as

$$\mu^{u_{[t,t_f]}}(B_t, B_x, B_u) := \mathbb{E}_{\mathcal{F}_t}^{u_{[t,t_f]}} \int_{B_t \cap [t,t_f]} \mathbb{I}_{B_x}(x_s) \cdot \mathbb{I}_{B_u}(u_s)\, \mathsf{d}s, \tag{15}$$

for arbitrary Borel sets $B_t \subset [t_0, t_f]$, $B_x \subset \mathbb{R}^n$, $B_u \subset U$, where $\mathbb{I}_B$ denotes the indicator function of the set B. We also define the *terminal state occupation measure* as

$$\kappa^{u_{[t,t_f]}}(B_x) := \mathsf{P}_{\mathcal{F}_t}^{u_{[t,t_f]}}\Big(x_{t_f} \in B_x\Big). \tag{16}$$

for an arbitrary Borel set $B_x \subset \mathbb{R}^n$.

Then for every twice continuously differentiable function $v \in C^2\left([t, t_f] \times \mathbb{R}^n\right)$

$$\left\langle v_{t_f}, \kappa^{u_{[t,t_f]}}\right\rangle - \left\langle \mathcal{A}v, \mu^{u_{[t,t_f]}}\right\rangle = v(t, x_t), \tag{17}$$

where

$$\left\langle v, \kappa^{u_{[t,t_f]}}\right\rangle = \int_{\mathbb{R}^n} v(t_f, x)\, \kappa^{u_{[t,t_f]}}(\mathsf{d}x), \tag{18}$$

$$\left\langle \mathcal{A}v, \mu^{u_{[t,t_f]}}\right\rangle = \int_{[t,t_f]\times\mathbb{R}^n\times U} \mathcal{A}^u v(s, x)\, \mu^{u_{[t,t_f]}}(\mathsf{d}s, \mathsf{d}x, \mathsf{d}u), \tag{19}$$

and $\mathcal{A}$ is the infinitesimal operator of the Markov process (1) and is given as

$$\mathcal{A}^u v(t,x) = \frac{\partial v(t,x)}{\partial t} + \left[\frac{\partial v(t,x)}{\partial x}\right]^\top f(t,x,u) + \frac{1}{2}\mathrm{tr}\left(g(t,x)^\top g(t,x)\, \frac{\partial^2 v(t,x)}{\partial x^2}\right). \tag{20}$$

Since the relation (17) holds for all $v \in C^2\left([t, t_f] \times \mathbb{R}^n\right)$, we can write it (see, e.g., [22], [26]) as

$$\kappa^{u_{[t,t_f]}} - \mathcal{A}^* \mu^{u_{[t,t_f]}} = \bar{\delta}_{(t,x_t)}, \tag{21}$$

where $\bar{\delta}_{(t,x_t)}$ is the Dirac measure, and $\mathcal{A}^*$ is the adjoint of (20) defined as the operator satisfying

$$\left\langle \mathcal{A}v, \mu\right\rangle = \left\langle v, \mathcal{A}^*\mu\right\rangle \tag{22}$$

for every Borel measure μ, and any twice continuously differentiable function $v \in C^2\left([t_0, t_f) \times \mathbb{R}^n\right)$.

Hence, by defining (see [22])

$$\mathcal{M}_{PB} := \Big\{(\mu, \kappa) \in \mathfrak{M}_+\left([t, t_f] \times \mathbb{R}^n \times U\right) \times \mathfrak{M}_+\left(\mathbb{R}^n\right) : \|\mu\| \leq t_f - t_0,\ \ \|\kappa\| \leq 1.\Big\}, \tag{23}$$

$$\mathcal{M}_{\mathcal{A}} := \Big\{(\mu, \kappa) \in \mathfrak{M}_\pm\left([t, t_f] \times \mathbb{R}^n \times U\right) \times \mathfrak{M}_\pm\left(\mathbb{R}^n\right) : \kappa - \mathcal{A}^*\mu = \bar{\delta}_{(t,x_t)}.\Big\}, \tag{24}$$

we can identify a subspace in the space of signed measures which tightly embeds the space of all occupation measures generated by all $u_{[t,t_f]} \in \mathcal{U}$, i.e.,

$$\mathcal{M}_{\mathcal{U}} := \left\{ \left(\mu^{u_{[t,t_f]}}, \kappa^{u_{[t,t_f]}}\right) : u_{[t,t_f]} \in \mathcal{U} \right\} \subset \mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A}} \tag{25}$$

In the remainder of this section, we provide measure theoretic equivalents of the constraints introduced in Section III and present (without proof, due to space limitations) the associated optimality conditions established by convex duality approaches (see [22], [26] for detailed derivations).

A. Almost Surely Satisfactions

The constraints (4) and (5) can be written as

$$\kappa^{u_{[\bullet,t_f]}} = \bar{\delta}_{(t_f,m_d)} \tag{26}$$

since, by the definition of the Dirac probability measure, for every Borel sets $B_t \subset [t_0, t_f]$, $B_x \subset \mathbb{R}^n$ we have

$$\bar{\delta}_{(t,x)}(B_t, B_x) = \begin{cases} 1, & \text{if } t \in B_t \text{ and } x \in B_x, \\ 0, & \text{otherwise.} \end{cases} \tag{27}$$

and, hence,

$$\mathsf{P}_{\mathcal{F}_\bullet}^{u_{[\bullet,t_f]}}\left(x_{t_f} \in B_x\right) = \int_{B_x} \bar{\delta}_{(t_f,m_d)}(\mathsf{d}x) \tag{28}$$

Since $\bar{\delta}_{(t_f,m_d)}$ can be regarded as a special case of a general desired probability distribution p_d, we delay the presentation of the optimal results until Section IV-C.

It is worth remarking that both of the above cases are likely to yield inexistence of a solution, i.e., such strong constraining the terminal state leads to ill-posedness of the associated optimal control problem.

B. Satisfaction in Expectation

While in [19] the optimality conditions for problems with expectation constraints are established in the form of the Terminally Constrained Stochastic Minimum Principle (TC-SMP), a separate (but closely related) set of optimality conditions are presented here based upon convex duality methods.

Using occupation measures, the constraints (6) and (7) can be expressed as $\langle x - m_d, \kappa \rangle = 0$. Accordingly, we define the weak value function as

$$W(t, x_t) := \min_{(\mu,\kappa) \in \mathcal{M}_{PB} \cap \mathcal{M}_{\mathcal{A},E}} \langle \ell, \mu \rangle + \langle L, \kappa \rangle, \tag{29}$$

with

$$\mathcal{M}_{\mathcal{A},E} := \Big\{ (\mu, \kappa) \in \mathfrak{M}_\pm\left([t, t_f] \times \mathbb{R}^n \times U\right) \times \mathfrak{M}_\pm\left(\mathbb{R}^n\right) : \kappa - \mathcal{A}^*\mu = \bar{\delta}_{(t,x_t)}, \ \langle x - m_d, \kappa \rangle = 0. \Big\}. \tag{30}$$

Then by invoking Rockafellar duality theorem (see [22], [26]) together with the equality of the weak and strong value functions, i.e., $W(t, x) = V(t, x)$ for all $(t, x) \in [t_0, t_f] \times \mathbb{R}^n$, due to the tightness of embedding [22], [26], we obtain the representation of the value function as

$$\begin{aligned} V(t, x_t) = \sup_{\beta_{t,x_t} \in \mathbb{R},\ v \in C^2([t,t_f] \times \mathbb{R}^n)} \Big\{ & v(t, x_t) \\ \text{s.t. } & \frac{\partial v(s,x)}{\partial s} + \left[\frac{\partial v(s,x)}{\partial x}\right]^\top f(s,x,u) \\ & + \frac{1}{2}\mathrm{tr}\Big(g(s,x)^\top g(s,x) \frac{\partial^2 v(s,x)}{\partial x^2}\Big) + \ell(s,x,u) \geq 0, \\ & \text{and } v(t_f, x) \leq L(x) + \beta_{t,x_t}(x - m_d), \\ & \text{for all } (s,x,u) \in [t, t_f] \times \mathbb{R}^n \times U \Big\}. \end{aligned} \tag{31}$$

It shall be remarked that the scalar $\beta_{t,x_t} \in \mathbb{R}$ is, in general, different under each filtration $\mathcal{F}_t$, i.e., for different values of $(t, x) \in [t, t_f] \times \mathbb{R}^n$. This, indeed, reflects the inherent *time-inconsistency* [34] in stochastic control problems with covariance constraints [35].

C. Satisfaction in Probability Distributions

1) Exact Assignment of Total Probability Distribution: It has been established in [22] that the value function for this case is identified by

$$\begin{aligned} V(t_0, x_0) = & \int_{\mathbb{R}^n} L(x)\mathsf{p}_d(\mathsf{d}x) \\ + \sup_{v \in C^2([t_0,t_f] \times \mathbb{R}^n)} \Big\{ & v(t_0, x_0) - \int_{\mathbb{R}^n} v(t_f, x)\mathsf{p}_d(\mathsf{d}x), \\ \text{s.t. } & \frac{\partial v(s,x)}{\partial s} + \left[\frac{\partial v(s,x)}{\partial x}\right]^\top f(s,x,u) \\ & + \frac{1}{2}\mathrm{tr}\Big(g(s,x)^\top g(s,x) \frac{\partial^2 v(s,x)}{\partial x^2}\Big) + \ell(s,x,u) \geq 0, \\ & \text{for all } (s,x,u) \in [t_0, t_f] \times \mathbb{R}^n \times U \Big\}. \end{aligned} \tag{32}$$

It is worth remarking that since in this class of problems, $\kappa^{u_{[t,t_f]}} = \mathsf{p}_d$ is fixed, then $\int_{\mathbb{R}^n} L(x)\kappa^{u_{[t,t_f]}}(\mathsf{d}x) = \int_{\mathbb{R}^n} L(x)\mathsf{p}_d(\mathsf{d}x)$ is a control-invariant constant in the problem and, hence, without loss of generality, we can assume $L(x) = 0$ and obtain the same representation as in [22, Theorem 4]. However, since other cases with non-fixed terminal distributions $\kappa^{u_{[t,t_f]}}$ are also studied in this paper, we present the results of this case with the constant term $\int_{\mathbb{R}^n} L(x)\mathsf{p}_d(\mathsf{d}x)$ included in the representation.

We can also use (32) to write the optimality conditions for Section IV-A by setting $\mathsf{p}_d = \bar{\delta}_{(t_f,m_d)}$. Due to limitations in space, these results are only displayed in Table I and not presented separately.

2,3) Exact Assignment of the Expectation(s) and Containment of Covariance(s): The constraints (11) and (13), as before, are expressed as $\langle x - m_d, \kappa \rangle = 0$. Moreover, the constraints (12) and (14) can be expressed by a set of inequality constraints[2] which we express by $\langle h_\Sigma(x), \kappa \rangle \leq 0$.

[2]Including the facts that if $\Sigma = [\sigma_{ij}] \preccurlyeq 0$ then $\sum \sigma_{ii} \geq 0$, $|\sigma_{ij}| \leq \sqrt{(|\sigma_{ii}\sigma_{jj}|)}$ and $|\sigma_{ij}| \leq \max_i \sigma_{ii}$.

Terminal Constraint	Optimization Parameters	Suprimizing Value	HJ Boundary Inequality
$\mathsf{P}^{u_{[t_0,t_f]}}_{\mathcal{F}_{t_0}}(x_{t_f}=m_d)=1$	$v\in C^2([t_0,t_f]\times\mathbb{R}^n)$	$v(t_0,x_0)-v(t_f,m_d)$	$v(t_f,x):$ free
$\mathsf{P}^{u_{[t,t_f]}}_{\mathcal{F}_{t}}(x_{t_f}=m_d)=1,\ t\in[t_0,t_f]$	$v\in C^2([t,t_f]\times\mathbb{R}^n)$	$v(t,x_t)-v(t_f,m_d)$	$v(t_f,x):$ free
$\mathbb{E}^{u_{[t_0,t_f]}}_{\mathcal{F}_{t_0}}[x_{t_f}]=m_d$	$\beta\in\mathbb{R}$ $v\in C^2([t_0,t_f]\times\mathbb{R}^n)$	$v(t_0,x_0)$	$v(t_f,x)\leq L(x)+\beta(x-m_d)$
$\mathbb{E}^{u_{[t,t_f]}}_{\mathcal{F}_{t}}[x_{t_f}]=m_d, \qquad t\in[t_0,t_f]$	$\beta_{t,x_t}\in\mathbb{R}$ $v\in C^2([t,t_f]\times\mathbb{R}^n)$	$v(t,x_t)$	$v(t_f,x)\leq L(x)+\beta_{t,x_t}(x-m_d)$
$\mathsf{P}^{u_{[t_0,t_f]}}_{\mathcal{F}_{t_0}}(x_{t_f}\in B_x)=\int_{B_x}\mathsf{p}_d(\mathsf{d}x)$	$v\in C^2([t_0,t_f]\times\mathbb{R}^n)$	$v(t_0,x_0)-\int_{\mathbb{R}^n}v(t_f,x)\mathsf{p}_d(\mathsf{d}x)$	$v(t_f,x):$ free
$\begin{cases}\mathbb{E}^{u_{[t_0,t_f]}}_{\mathcal{F}_{t_0}}[x_{t_f}]=m_d,\\ \mathsf{cov}^{u_{[t_0,t_f]}}_{\mathcal{F}_{t_0}}(x_{t_f})\preccurlyeq\Sigma_d\end{cases}$	$\beta\in\mathbb{R}$ $\Gamma\in\mathbb{R}^{n\times n}_{\succeq 0}$ $v\in C^2([t_0,t_f]\times\mathbb{R}^n)$	$v(t_0,x_0)$	$v(t_f,x)\leq L(x)+\beta(x-m_d)$ $+\frac{1}{2}(x-m_d)^\top\Gamma(x-m_d)$
$\begin{cases}\mathbb{E}^{u_{[t,t_f]}}_{\mathcal{F}_{t}}[x_{t_f}]=m_d,\\ \mathsf{cov}^{u_{[t,t_f]}}_{\mathcal{F}_{t}}(x_{t_f})\preccurlyeq\Sigma_d,\end{cases} \quad t\in[t_0,t_f]$	$\beta_{t,x_t}\in\mathbb{R}$ $\Gamma_{t,x_t}\in\mathbb{R}^{n\times n}_{\succeq 0}$ $v\in C^2([t,t_f]\times\mathbb{R}^n)$	$v(t,x_t)$	$v(t_f,x)\leq L(x)+\beta_{t,x_t}(x-m_d)$ $+\frac{1}{2}(x-m_d)^\top\Gamma_{t,x_t}(x-m_d)$

TABLE I: Summary of the results of the Distributionally Constrained Convex Duality Optimal Control (DC-CDOC) in the form of $\sup_{\text{Optimization Parameters}}\{\text{Suprimizing Value} \mid \mathcal{A}v+\ell\geq 0\wedge \text{HJ Boundary Inequality}\}$ corresponding to different forms terminal state constraints.

According, we define

$$\mathcal{M}_{PB\Sigma}:=\Big\{(\mu,\kappa)\in\mathfrak{M}_+\left([t,t_f]\times\mathbb{R}^n\times U\right)\times\mathfrak{M}_+\left(\mathbb{R}^n\right): \Big\langle h_\Sigma(x),\kappa\Big\rangle\leq 0,\ \ \|\mu\|\leq t_f-t_0,\ \ \|\kappa\|\leq 1.\Big\}, \quad (33)$$

and write the value function as

$$V(t,x_t):=\min_{(\mu,\kappa)\in\mathcal{M}_{PB\Sigma}\cap\mathcal{M}_{\mathcal{A},E}}\Big\langle \ell,\mu\Big\rangle+\Big\langle L,\kappa\Big\rangle, \qquad (34)$$

Then by invoking Rockafellar duality theorem we obtain the representation of the value function as

$$\begin{aligned}V(t,x_t)=&\sup_{\beta_{t,x_t}\in\mathbb{R},\ \Gamma_{t,x_t}\in\mathbb{R}^{n\times n}_{\succeq 0},\ v\in C^2([t,t_f]\times\mathbb{R}^n)}\Bigg\{v(t,x_t)\\ \text{s.t. }&\frac{\partial v(s,x)}{\partial s}+\left[\frac{\partial v(s,x)}{\partial x}\right]^\top f(s,x,u)\\ &+\frac{1}{2}\mathrm{tr}\Big(g(s,x)^\top g(s,x)\,\frac{\partial^2 v(s,x)}{\partial x^2}\Big)+\ell(s,x,u)\geq 0,\\ &\text{and } v(t_f,x)\leq L(x)+\beta_{t,x_t}(x-m_d)\\ &\qquad+\frac{1}{2}(x-m_d)^\top\Gamma_{t,x_t}(x-m_d),\\ &\text{for all } (s,x,u)\in[t,t_f]\times\mathbb{R}^n\times U\Bigg\}. \end{aligned} \qquad (35)$$

V. Conceptual Algorithm

Based on the results of Section IV summarized in Table I, one can employ the following conceptual algorithm to establish the value function of each of the constrained stochastic optimal control problems.

Step 0: Set the iteration counter to $k=0$.

Step 1: If the corresponding HJ boundary inequality in Table I is free, then initiate the algorithm with an arbitrary terminal cost function $L^k(x)$; otherwise, initiate $L^k(x)$ by selecting arbitrary values for β^k (and Γ^k as appropriate).

Step 2: Solve the HJB equation[3]

$$\frac{\partial v^k(t,x)}{\partial t}+\min_{u\in U}\left\{\left(\frac{\partial v^k(t,x)}{\partial x}\right)^\top f(x,u,t)+\ell(x,u,t)\right\}=0, \qquad (36)$$

subject to $v^k(t_f,x)=L^k(x)$.

Step 3: Evaluate $v^k(t_0,x_0)-\int_{\mathbb{R}^n}v^k(t_f,x)\mathsf{p}_d(\mathsf{d}x)$.

Step 4: Update $L^{k+1}(x)$ using an ascent direction[4] for the cost in the column "Suprimizing Value" in Table I.

VI. Concluding Remarks

The Distributionally Constrained Convex Duality Optimal Control (DC-CDOC) is a powerful tool for identifying the value function in a large class of nonlinear stochastic optimal control problems subject to a wide range of Gaussian and

[3]If a classical solution does not exist, one needs to consider an additional supremization over subsolutions of the HJB, i.e., the family of functions (indexed by another iteration j, satisfying the HJ inequalities $\frac{\partial v_j^k(t,x)}{\partial t}+\min_{u\in U}\left\{\left(\frac{\partial v_j^k(t,x)}{\partial x}\right)^\top f(x,u,t)+\ell(x,u,t)\right\}\geq 0$, but subject to the equality conditions $v_j^k(t_f,x)=L^k(x)$ for all j. However, it can be shown that the suprimizing function (over all j) converges to the viscosity solution of the HJB equation (36).

[4]Due to the computationally expensive nature of the cost, and the infinite dimensionality of the space of terminal costs, derivative-free methods such as Bayesian optimization shall be used in this general procedure.

non-Gaussian terminal distribution constraints under various filtrations. Despite the inherent time-inconsistency, due to the involvement of the second moment (covariance) and higher order moments of the state distribution (which precludes the implementation of Bellman's principle of optimality [36] to arrive at a Hamilton-Jacobi-Bellman (HJB) equation) the DC-CDOC framework effectively identifies the value function via an optimization problem over a class of Hamilton-Jacobi (HJ) inequalities. A key advantage of DC-CDOC is that the identification is performed through a family of smooth (sub-)solutions despite the potential nonsmoothness of the original value function. This opens the door for the development of new numerical algorithms by constructing parameterized families of smooth functions and constructing an optimization problem with parameters described in Table I. However, special care must be taken as (i) the sequence of test functions might converge to a value smaller than the optimal cost which suggests that the family of functions does not contain a function characterizing the value function, or (ii) a maximum might not exist; therefore, the supremum (and thus the value function) must be characterized by examining the limiting behavior of the associated family of functions.

ACKNOWLEDGMENT

Discussions concerning this work with Panagiotis Tsiotras are gratefully acknowledged.

REFERENCES

[1] A. Hotz and R. E. Skelton, "Covariance Control Theory," *International Journal of Control*, vol. 46, no. 1, pp. 13–32, 1987.

[2] J.-H. Xu and R. Skelton, "An Improved Covariance Assignment Theory for Discrete Systems," *IEEE transactions on Automatic Control*, vol. 37, no. 10, pp. 1588–1591, 1992.

[3] K. Yasuda, R. E. Skelton, and K. M. Grigoriadis, "Covariance Controllers: A New Parametrization of the Class of all Stabilizing Controllers," *Automatica*, vol. 29, no. 3, pp. 785–788, 1993.

[4] K. M. Grigoriadis and R. E. Skelton, "Minimum-Energy Covariance Controllers," *Automatica*, vol. 33, no. 4, pp. 569–578, 1997.

[5] Y. Chen, T. T. Georgiou, and M. Pavon, "Optimal Steering of a Linear Stochastic System to a Final Probability Distribution, Part I," *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1158–1169, 2015.

[6] ——, "Optimal Steering of a Linear Stochastic System to a Final Probability Distribution, Part II," *IEEE Transactions on Automatic Control*, vol. 61, no. 5, pp. 1170–1180, 2015.

[7] A. Beghi, "On the Relative Entropy of Discrete-Time Markov Processes with Given End-Point Densities," *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1529–1535, 1996.

[8] E. Bakolas, "Optimal covariance control for discrete-time stochastic linear systems subject to constraints," in *Proceedings of the 55th IEEE Conference on Decision and Control*, 2016, pp. 1153–1158.

[9] ——, "Finite-Horizon Covariance Control for Discrete-Time Stochastic Linear Systems subject to Input Constraints," *Automatica*, vol. 91, pp. 61–68, 2018.

[10] M. Goldshtein and P. Tsiotras, "Finite-Horizon Covariance Control of Linear Time-Varying Systems," in *Proceedings of the 56th IEEE Conference on Decision and Control*, 2017, pp. 3606–3611.

[11] K. Okamoto, M. Goldshtein, and P. Tsiotras, "Optimal Covariance Control for Stochastic Systems Under Chance Constraints," *IEEE Control Systems Letters*, vol. 2, no. 2, pp. 266–271, 2018.

[12] K. Okamoto and P. Tsiotras, "Optimal Stochastic Vehicle Path Planning using Covariance Steering," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2276–2281, 2019.

[13] A. Halder and E. Wendel, "Finite Horizon Linear Quadratic Gaussian Density Regulator with Wasserstein Terminal Cost," in *Proceedings of the American Control Conference*, 2016, pp. 7249–7254.

[14] K. F. Caluya and A. Halder, "Finite Horizon Density Control for Static State Feedback Linearizable Systems," in *Proceedings of the 2020 American Control Conference (ACC)*. IEEE, 2020, pp. 3577–3582.

[15] ——, "Wasserstein proximal algorithms for the schrödinger bridge problem: Density control with nonlinear drift," *IEEE Transactions on Automatic Control*, vol. 67, no. 3, pp. 1163–1178, 2021.

[16] I. Nodozi and A. Halder, "Schrödinger meets kuramoto via feynman-kac: Minimum effort distribution steering for noisy nonuniform kuramoto oscillators," in *2022 IEEE 61st Conference on Decision and Control (CDC)*. IEEE, 2022, pp. 2953–2960.

[17] J. Ridderhof, K. Okamoto, and P. Tsiotras, "Nonlinear Uncertainty Control with Iterative Covariance Steering," in *2019 IEEE 58th Conference on Decision and Control (CDC)*. IEEE, 2019, pp. 3484–3490.

[18] Z. Yi, Z. Cao, E. Theodorou, and Y. Chen, "Nonlinear covariance control via differential dynamic programming," in *2020 American Control Conference (ACC)*. IEEE, 2020, pp. 3571–3576.

[19] A. Pakniyat and P. Tsiotras, "Steering the State of Linear Stochastic Systems: A Constrained Minimum Principle Formulation," in *Proceedings of the 2021 IEEE American Control Conference (ACC)*, 2021, pp. 1300–1305.

[20] ——, "Partially Observed Steering the State of Linear Stochastic Systems," in *Proceedings of the 60th IEEE Conference on Decision and Control (CDC)*, 2021, pp. 3780–3785.

[21] A. Pakniyat, "Theoretical Guarantees for Satisfaction of Terminal State Constraints for Nonlinear Stochastic Systems," in *IUTAM Symposium on Optimal Guidance and Control for Autonomous Systems*. Springer, 2023, pp. 135–163.

[22] ——, "A Convex Duality Approach for Assigning Probability Distributions to the State of Nonlinear Stochastic Systems," *IEEE Control Systems Letters*, vol. 6, pp. 3080–3085, 2022.

[23] R. B. Vinter and R. M. Lewis, "The Equivalence of Strong and Weak Formulations for Certain Problems in Optimal Control," *SIAM Journal on Control and Optimization*, vol. 16, no. 4, pp. 546–570, 1978.

[24] ——, "A Necessary and Sufficient Condition for Optimality of Dynamic Programming Type, Making No a Priori Assumptions on the Controls," *SIAM Journal on Control and Optimization*, vol. 16, no. 4, pp. 571–583, 1978.

[25] D. Vermes, "Optimal Control of Piecewise Deterministic Markov Processes," *Stochastics: An International Journal of Probability and Stochastic Processes*, vol. 14, no. 3, pp. 165–207, 1985.

[26] W. H. Fleming and D. Vermes, "Convex Duality Approach to the Optimal Control of Diffusions," *SIAM journal on control and optimization*, vol. 27, no. 5, pp. 1136–1155, 1989.

[27] H. Zhu, "Convex Duality for Finite-Fuel Problems in Singular Stochastic Control," *Journal of optimization theory and applications*, vol. 75, no. 1, pp. 155–181, 1992.

[28] A. G. Bhatt and V. S. Borkar, "Occupation Measures for Controlled Markov Processes: Characterization and Optimality," *The Annals of Probability*, pp. 1531–1562, 1996.

[29] T. Kurtz and R. Stockbridge, "Existence of Markov Controls and Characterization of Optimal Markov Controls," *SIAM Journal on Control and Optimization*, vol. 36, no. 2, pp. 609–653, 1998.

[30] M. Cho and R. Stockbridge, "Linear Programming Formulation for Optimal Stopping Problems," *SIAM Journal on Control and Optimization*, vol. 40, no. 6, pp. 1965–1982, 2002.

[31] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, "Nonlinear Optimal Control via Occupation Measures and LMI-Relaxations," *SIAM journal on control and optimization*, vol. 47, no. 4, pp. 1643–1666, 2008.

[32] E. Pauwels, D. Henrion, and J.-B. Lasserre, "Linear conic optimization for inverse optimal control," *SIAM Journal on Control and Optimization*, vol. 54, no. 3, pp. 1798–1825, 2016.

[33] Y. Chen and A. D. Ames, "Duality between Density Function and Value Function with Applications in Constrained Optimal Control and Markov Decision Process," *arXiv preprint arXiv:1902.09583*, 2019.

[34] R. H. Strotz, *Myopia and Inconsistency in Dynamic Utility Maximization*. Springer, 1973.

[35] T. Björk and A. Murgoci, "A General Theory of Markovian Time Inconsistent Stochastic Control Problems," *SSRN Electronic Journal, 1694759*, 2010.

[36] R. Bellman, "Dynamic Programming," *Science*, vol. 153, no. 3731, pp. 34–37, 1966.